IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

First Named Inventor : Robert L. Chambers	
Appln. No.: 10/799,356	Confirmation No.: 6338
Filed : March 12, 2004	Group Art Unit: 2626
For : SPEECH RECOGNITION USING CATEGORIES AND SPEECH PREFIXING	Examiner: Daniel Demelash Abebe
Docket No.: M61.12-0601	

BRIEF FOR APPELLANTS

Mail Stop Appeal Brief-Patents Commissioner for Patents P.O. Box 1450 Alexandria, VA 22313-1450	**Electronically Filed on July 21, 2008**

Sir:

This is an appeal from a Final Office Action mailed November 19, 2007 in which claims 1-15 were rejected. Appellants respectfully submit that claims 1-15 are allowable, and request that the Board reverse the rejection of claims 1-15 and find that claims 1-15 are in condition for allowance.

Contents

REAL PARTY IN INTEREST

Microsoft Corporation, a corporation organized under the laws of the state of Washington, and having a place of business at One Microsoft Way, Redmond, WA, 98052, has acquired the entire right, title and interest in and to the invention, the application, and any and all patents to be obtained therefor, as set forth in the Assignment filed with the patent application and recorded on Reel 015424, Frame 0621.

NO RELATED APPEALS OR INTERFERENCES

There are no known related appeals or interferences which will directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

STATUS OF THE CLAIMS

Claims 1-15 were originally presented and have not been amended. Thus, the pending and rejected claims 1-15 are the subject of the present appeal.

STATUS OF AMENDMENTS

No amendments have been filed subsequent to the mailing of the Final Office Action.

SUMMARY OF CLAIMED SUBJECT MATTER

1.   **Introduction**

The present invention relates to computerized speech recognition. More particularly, the present invention relates to an apparatus and methods to remove ambiguity with respect to a speech recognition system.

2.   Brief Background

Speech recognition is a technology that has a number of useful applications that allow people to interface with computing systems using their voices. These applications include: allowing a user to dictate text into a document; allowing a user to issue commands to one or more computer systems via speech; improving automated telephony systems; and many other applications. Such systems are useful in large centralized-server applications, such as computerized telephony processing systems; user interaction with desktop computing products; and even improved interaction and control of mobile computing devices.

Speech recognition is known and is being actively researched as perhaps the future of human interaction with computing devices. While speech recognition technology has progressed rapidly, it has not been perfected. Currently, speech recognition requires substantial computing resources and has not achieved 100% recognition accuracy. This is partly due to inherent ambiguities in human language, and also due, in part, to varying domains over which user speech may be applied.

In a speech recognition system supporting multiple third party applications, grammars from different applications will often accept the same speech command. Thus, there is inherent ambiguity in which application should execute the command when a user issues such a command. For example, the command "send message" may be accepted by grammars from both Microsoft Outlook and Microsoft Messenger, both of which are available from Microsoft Corporation of Redmond, Washington. In addition to the ambiguity created by determining which application to direct a command to, there is reduced recognition accuracy when a command of a much more constrained grammar (such as would be directed to a specific application), is required to be recognized by a larger grammar such as a large vocabulary dictation grammar.

Providing a speech recognition system and methods that could facilitate the direction of user speech to specific programs and/or modules as well as attempt to recognize such speech with specifiable grammars would represent an improvement to speech recognition without adding significant complexity to the user experience.

3.   The Present Invention

Claims 1, 12 and 13 are the only independent claims on appeal.

Claim 1 provides a method of recognizing speech. The method includes detecting a predefined prefix (as illustrated at reference numeral 604 in Fig. 6, and described, at least, on page 16, line 29 – page 17, line 2). Text from speech following the prefix is recognized using a set of grammars associated with the detected prefix (as illustrated in Fig. 6 at reference numeral 608 and described, at least, on page 17, lines 2-4). The recognized text is directed to a target associated with the set of grammars (as illustrated at reference numeral 610 in Fig. 6, and described, at least, on page 17, lines 4-7).

Claim 12 provides a data structure (as illustrated at reference numeral 400 in Fig. 4, and described, at least, on page 14, line 29 – page 15, line 23) for storing information relative to a speech recognition category. The data structure includes a prefix field (as illustrated at reference numeral 404, and described, at least, on page 15, lines 7 – 10), a grammar field (as illustrated at reference numeral 408, and described, at least, on page 15, lines 12 – 14), an IsRequired field (as illustrated at reference numeral 410, and described, at least, on page 15, lines 14 – 16) and a parent field (as illustrated at reference numeral 402, and described, at least, on page 15, lines 2 – 7). The prefix field stores a prefix, which when recognized will cause the category to become an ActiveCategory. The grammar field indicates a set of grammars to use with the category. The IsRequired field indicates whether the prefix must be uttered in order to direct speech to the set of grammars. The parent field indicates a parent category.

Claim 13 provides a speech recognition system. The system includes an input for receiving speech (as illustrated at reference numeral 163 in Fig. 1, and described, at least, on page 8, lines 24 – 27). A processor (as illustrated at reference numeral 120 in Fig. 1 and reference numeral 202 in Fig. 2) recognizes speech using a set of one or more grammars (as illustrated at reference numeral 500 in Fig. 5, and described, at least, on page 15, lines 24-30). The processor is adapted to recognize a prefix (as illustrated at reference numeral 604 in Fig. 6, and described, at least, on page 16, line 29 – page 17, line 2) associated with a desired category, and to

recognize speech using the set of one or more grammars associated with the desired category when the prefix is recognized to generate an output (as illustrated in Fig. 6 at reference numeral 608 and described, at least, on page 17, lines 2-4). The output is provided to a target associated with the recognized grammar (as illustrated at reference numeral 610 in Fig. 6, and described, at least, on page 17, lines 4-7).

GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Whether claims 1-15 are allowable over Gould et al. (U.S. Pat. No. 6,839,669 – hereinafter "Gould").

Appellants respectfully submit that claims 1-15 are patentable over Gould and request that the Board find likewise and accordingly reverse the rejection of claims 1-15 and find these claims allowable.

ARGUMENT

1.   Introduction: Claims 1-15 Should Be Allowed

With this appeal, Appellants respectfully request that the Board reverse the rejection of claims 1-15.

2.   Anticipation

As set forth in Section Three of the Final Office Action, 35 U.S.C. §102(e) provides that a person shall be entitled to a patent unless -

> "(e) the invention was described in ... a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent..."

Further, Appellants respectfully note that the Federal Circuit has provided guidance with respect to anticipation. The Federal Circuit has held anticipation to be present, "If

every limitation in a claim is found in a single prior art reference." See Nystrom v. Trex Co., 71 U.S.P.Q.2.d. 1241 (Fed. Cir. 2004).

2.1   Claim 1 is Not Anticipated by Gould

Section Four of the Office Action indicated that independent claim 1, among others, was rejected under 35 U.S.C. §102(e) as being anticipated by Gould. Independent claim 1 provides,

> A method of recognizing speech, the method comprising:
> detecting a predefined prefix;
> recognizing text from speech following the prefix using a set of grammars associated with the detected prefix; and
> directing recognized text to a target associated with the set of grammars.

Gould is a lengthy reference directed to performing actions identified in recognized speech. Generally, the Gould reference provides techniques that allow users to draw benefits of large vocabulary continuous speech recognition while away from the desktop. See column 1, lines 59-62. More specifically, the techniques generally employ a voice recorder (See column 2, line 9) or direct connection via a telephone (See column 2, line 27) where the user issues a voice command. "After transferring data from the recorder, the interface software 380 provides the digital samples for an action item to the speech recognition software 360." Column 7, lines 30-32. Further, Gould provides, "In general, the speech recognition software analyzes the digital samples to produce a sequence of text, and provides this sequence to the interface software 380." See column 7, lines 35-38. Columns 7-65 of the Gould reference are directed generally to the manner in which the text provided by the speech recognition engine is parsed to discern commands and action items embedded therein. Only beginning at Column 65, Line 52 does Gould actually discuss mechanics of speech recognition.

Section Four of the Final Office Action alleged that the feature of detecting a pre-defined prefix is provided by the command notation of Gould, "computer please" recited in column 8, lines 38-59; column 33, lines 24-35; and column 42, lines 43-45. However, the entire disclosure of Gould in Column 7 through column 65, line 50, is directed, not to the recognition

of speech, but to the characterization of the recognized text. Thus, the "computer please" utterance is first transformed to text by speech recognition software 360, and then the actual text is recognized as a command. This is an important distinction because the manner in which the speech recognition itself is performed is not being affected by the utterance. In distinct contrast, independent claim 1 features a method of <u>recognizing speech</u> that includes detecting a predefined prefix. Then, <u>speech</u> following the prefix <u>is recognized</u> using a set of grammars <u>associated with the detected prefix</u>. While column 66, lines 66-67 of Gould provides, "different constraint grammars may be active at different times…" there is no teaching or suggestion in the entire encyclopedic disclosure of Gould of a user's utterance actually having any ability to select, or otherwise determine, any constraint grammars to be applied to speech following a detected prefix. As set forth on page 12 of the Appellants' specification, "Grammar categories can be deterministically selected by uttering user-specifiable speech prefixes." Further, "In this manner, the word will be recognized based upon a much more constrained grammar and recognition accuracy will be improved."

In Section Two of the Final Office Action, the Response to Arguments section asserted that the distinction between speech recognition and text characterization was not given patentable weight because the recitation occurs in Appellant's preamble. Appellants respectfully submit that this was erroneous for at least two reasons.

First, there is no *per se* rule that preamble recitations cannot be given patentable weight. Section 2111.02 of The Manual of Patent Examining Procedure (MPEP) cites Federal Circuit case law when it notes that "[t]he determination of whether a preamble limits a claim is made on a case-by-case basis in light of the facts in each case; there is no litmus test defining when a preamble limits the scope of a claim." Additionally, Appellants note that independent claim 1 is a method claim, and that the preamble recites a "method of recognizing speech." The MPEP cites a Court of Customs and Patent Appeals case when it notes, "[i]n considering the effect of the preamble in a claim directed to a method of treating or preventing pernicious anemia in humans by administering a certain vitamin preparation to 'a human in need thereof,' the court held that the claims' recitation of a patient or a human 'in need' gives life and meaning to the

preamble's statement of purpose." Citing <u>Kropa v. Robie</u>, 88 U.S.P.Q. 478, 481 (CCPA 1951). Similarly, Appellants respectfully submit that "recognizing speech" in the preamble of claim 1 gives life and meaning to the preamble.

The second reason why Appellants believe that the Final Office Action's refusal to give patentable weight to the speech recognition/text characterization distinction is that the distinction is also present in <u>the body of claim 1</u>. Specifically, claim 1, includes "recognizing text from speech following the prefix following the prefix using a set of grammars associated with the detected prefix." Recognizing text from speech that follows the prefix is speech recognition. Accordingly, Appellants respectfully believe that the Final Office Action conveniently ignored an important distinction between claim 1 and the Gould reference by erroneously characterizing the entire distinction as existing in the preamble and then erroneously asserting that it can be ignored since it is only in the preamble.

Notwithstanding that important distinction, claim 1 still recites yet another feature that is neither taught nor suggested by Gould. Specifically, the text that is recognized by the associated set of grammars is then <u>directed to a target associated with that set of grammars</u>. While Gould provides extensive parsing and processing of "recognized" text, there is no indication that recognized speech is provided to any target other than the speech recognition engine itself, or interface software 380/2320. Therefore, Appellants respectfully submit that original independent claim 1 is allowable over Gould. Further, Appellants respectfully submit that dependent claims 2-11 are similarly allowable by virtue of their dependency, either directly or indirectly from allowable claim 1.

2.2 Claim 12 is Not Anticipated by Gould

Section Four of the Final Office Action also indicated that independent claim 12 was rejected under 35 U.S.C. 102(e) as being anticipated by Gould. Independent claim 12 recites a data structure for storing information relative to a speech recognition category. The data structure includes a prefix field, a grammar field, an IsRequired field, and a parent field. As set forth above, Appellants respectfully submit that Gould simply does not teach prefixes in the context of speech

recognition, but instead identifies commands within detected text. Accordingly, Gould simply does not provide a prefix field. However, even if such a prefix field could be considered present in the Gould reference, there is no indication of a data structure providing a grammar field indicating a set of grammars to use with the category. While the Final Office Action referred to column 22, lines 1-14 of Gould, that portion of the Gould reference, and essentially the entire disclosure between column 7 and line 51 of column 65 is directed to text. The only discussion of constraint grammars begins after column 65, and there is simply no teaching or suggestion of selectable grammars associated with categories as set forth in independent claim 12. Further still, Appellants respectfully submit that even if Gould did provide prefixes for speech recognition and selectable grammars based on prefix detection, that there is still no data structure provided that meets each and every limitation of independent claim 12. Accordingly, Appellants respectfully submit that independent claim 12 is allowable over Gould.

2.3   Claim 13 is Not Anticipated by Gould

Section Four of the Final Office Action also indicated that independent claim 13, among others, was rejected under 35 U.S.C 102(e) as being anticipated by Gould. Claim 13 provides:

> A speech recognition system comprising:
> an input for receiving speech;
> a processor for recognizing speech using a set of one or more grammars;
> wherein the processor is adapted to recognize a prefix associated with a desired
>   category, and to recognize speech using the set of one or more grammars
>   associated with the desired category when the prefix is recognized to
>   generate an output; and
> wherein the output is provided to a target associated with the recognized grammar.

Claim 13 is directed towards a speech recognition system. The processor is used for recognizing speech using a set of one or more grammars. The processor, however, is adapted to recognize a prefix associated with the desired category and to recognize speech using the set of one or more grammars associated with the desired category when the prefix is recognized to generate an output. It is respectfully submitted that the Gould reference does not teach the

limitation from claim 13 "to recognize speech using the set of one or more grammars associated with the desired category when the prefix is recognized to generate an output." The prefix of claim 13 is associated with a desired category, which is, in turn, associated with a set of one or more grammars which can then be used to recognize speech. The Final Office Action asserted that the limitation "wherein the processor is adapted to recognize a prefix associated with a desired category" is met by Gould at column 31, line 1 - column 32, line 34 in the discussion of the "preamble." However, Appellants respectfully note, that the preamble "take a note" listed in line 32 is text that has been previously recognized. There is no indication that recognition of the text "take a note" while parsing recognized text somehow affects speech recognition. The preamble, which is cited by the Final Office Action, categorizes what the user is doing such as "the user is creating a task" (Column 31, Para. 2). However, this preamble-category relationship does not relate in any way to the grammar constraints of Gould. The grammar constraints of Gould are referred to as template or restrictions and may limit the words that may correspond to an utterance (col. 66, line 52-67). There is no relation drawn from the preamble to these constraints. It is therefore submitted that for at least these reasons that claim 13 is allowable over Gould. Moreover, Appellants respectfully submit that claims 14 and 15 are allowable as well by virtue of their dependency, either directly or indirectly, from allowable independent claim 13.

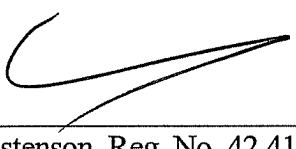3.   **Conclusion: Claims 1-15 should be allowed.**

In conclusion, Appellants respectfully submit that the rejection of claims 1-15 is improper, and that all claims 1-15 are in condition for allowance. Accordingly, Appellants respectfully request that the Board reverse the rejection of claims 1-15 and find that such claims are allowable.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By:_____
      Christopher R. Christenson, Reg. No. 42,413
      Suite 1400, 900 Second Avenue South
      Minneapolis, Minnesota 55402-3319
      Phone:(612) 334-3222 Fax:(612) 334-3312

Appendix A: Claims On Appeal

<u>Claims on appeal, as they currently stand:</u>

1.   A method of recognizing speech, the method comprising:

detecting a predefined prefix;

recognizing text from speech following the prefix using a set of grammars
            associated with the detected prefix; and

directing recognized text to a target associated with the set of grammars.

2.   The method of claim 1, wherein the set of grammars is specific to a computer
program.

3.   The method of claim 2, wherein the target is the computer program.

4.   The method of claim 3, wherein the prefix is at least one word related to the
computer program.

5.   The method of claim 1, wherein the prefix is a word.

6.   The method of claim 1, wherein the prefix is a phrase.

7.   The method of claim 1, wherein the predefined prefix, and set of grammars are
specified in a category.

8.   The method of claim 7, wherein the category also includes a field to identify a
parent of the category.

9.   The method of claim 7, wherein the category also includes a field to determine if a prefix is required to invoke the category.

10.   The method of claim 1, executed upon a desktop computer.

11.   The method of claim 1, executed upon a mobile computing device.

12.   A data structure for storing information relative to a speech recognition category, the data structure comprising:

> a prefix field, which when recognized will cause the category to become an ActiveCategory;
>
> a grammar field indicating a set of grammars to use with the category;
>
> an IsRequired field to indicate whether the prefix must be uttered in order to direct speech to the set of grammars; and
>
> a parent field indicating a parent category.

13.   A speech recognition system comprising:

> an input for receiving speech;
>
> a processor for recognizing speech using a set of one or more grammars;
>
> wherein the processor is adapted to recognize a prefix associated with a desired category, and to recognize speech using the set of one or more grammars associated with the desired category when the prefix is recognized to generate an output; and
>
> wherein the output is provided to a target associated with the recognized grammar.

14.   The recognizer of claim 13, wherein the prefix must be recognized in order activate the set of grammars.

15.　　　　The recognizer of claim 14, wherein the set of grammars is specific to a computer application.

Appendix B: Evidence

(None)

Appendix C: Related Proceedings

(None)